BEBR

FACULTY WORKING
PAPER NO. 1424

A Model of Efficiency Wages as a Signal of Firm Value

Lanny Arvan
Hadi S. Esfahani

WORKING PAPER SERIES ON THE POLITICAL ECONOMY OF INSTITUTIONS
NO. 5

College of Commerce and Business Administration
Bureau of Economic and Business Research
University of Illinois, Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 1424

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

December 1987

WORKING PAPER SERIES ON THE POLITICAL ECONOMY OF INSTITUTIONS NO. 5

A Model of Efficiency Wages as a Signal of Firm Value

Lanny Arvan, Associate Professor
Department of Economics

Hadi S. Esfahani, Assistant Professor
Department of Economics

We would like to thank the participants of the Economic Theory Workshop at the University of Illinois at Urbana-Champaign for several helpful suggestions, many of which found their way into the paper.

Abstract

Worker-posted performance bonds are a potential solution to the agency problem which arises when worker effort is not perfectly observable by the firm. Yet the explicit posting of performance bonds is rarely observed. Since the firm should prefer a contractual arrangement which entails performance bonds to one which relies on efficiency wages, because the bonding contract yields lower unit labor costs, an explanation for the lack of bonds in actuality is required. We propose such an explanation based on the idea that the implicit contract must resolve the firm moral hazard concerning bond expropriation. We model this moral hazard as a loan default problem. When the value of the firm is less than the value of the posted performance bonds, the firm will expropriate the bonds and go out of business. Workers will willingly post performance bonds when thy know that this inequality is reversed. In an environment where the firm is subject to a series of random shocks which are not observed by its workers, these workers will be unable to determine the value of the firm. We demonstrate that in such an environment the optimal implicit contract may require the firm to pay wages to workers in excess of their reservation wage, when labor productivity and thus labor demand is high. These efficiency wages act both as a signal to workers that there is no threat of bond expropriation and as an incentive device to elicit effort.

I. Introduction

Worker-posted performance bonds appear to be a viable solution to the agency problem which arises when worker effort is not perfectly observable by the firm and, consequently, when the firm must generate incentives for its workers to put forth effort. Such bonding contracts, if feasible, obviate the need for the firm to pay its workers in excess of their reservation wage, as suggested by the recently developed efficiency wage theory.[1] However, the explicit posting of performance bonds is rarely observed. Moreover, recent work by Abraham and Farber (1987) suggests that even implicit bonding arrangements, in the form of upward sloping wage-earnings profiles steeper than their concomitant productivity profiles, are less significant than was previously suspected.[2] Thus, one is led to seek explanations for why bonding does not occur more often.

The main theoretical argument that has been advanced to date as an explanation for this puzzle is that workers are liquidity constrained and, thus, unable to put up the requisite bond. Although this explanation may be plausible in certain cases, its general applicability faces severe limitations, because the explanation appears to require that when workers are somewhat liquid the firm will require these workers to post all available cash as performance bond, so that worker compensation need not be as high as it would be were there exclusive reliance on efficiency wages. This implies that the firm should rank its job applicants, who are otherwise similar, on the basis of their relative liquidity. Moreover, to the extent that the firm can perceive liquidity differences among its existing workforce, internal wage differentials should arise. We believe that this is the wrong line of approach to explain the lack of bonding in actuality, though liquidity constraints may bind in many instances.

---

[1] See, for example, Shapiro and Stiglitz (1984) and Bulow and Summers (1986). Yellen (1984), Akerlof and Yellen (1986), and Katz (1986) provide interesting reviews of the efficiency wage literature.

[2] See, for example, Medoff and Abraham (1980).

Our preferred explanation takes as its basis the firm moral hazard concerning bond expropriation. In the efficiency wage literature, this moral hazard is usually described as the firm deliberately mislabeling its workers as shirkers. It has been argued that this moral hazard is solved through the firm's desire to maintain its reputation in the labor market.[3, 4] The argument is that upon reneging, recruiting costs go up sufficiently to more than offset the gains resulting from bond expropriation. In a recent paper, we argue that bonding will dominate efficiency wages, when such reputational effects are viable.[5] However, the resulting equilibrium does not yield the first best solution. Although the labor market clears in this equilibrium, firms must earn reputational rents which are themselves distortionary, when compared to the solution where the firm moral hazard is absent. In our current paper, we assume instantaneous and certain detection of bond expropriation, in order to simplify the model and avoid extraneous issues, particularly the question of how detection of the firm moral hazard occurs.[6] Then we view the firm moral hazard as a loan default problem. That is, when the firm expropriates bond it takes the bond of all its workers simultaneously.[7] We also assume that the firm ceases operation once

---

[3] See, for example, Akerlof and Yellen (1986).

[4] Several authors have suggested mechanisms other than reputation to resolve the firm moral hazard problem in contracts with bonding. For example, Carmichael (1985) argues that the problem can be eliminated through the use of third party repositories for the performance bonds and through the designation of recipients of performance bonds that are appropriated in the event that some workers are detected shirking, other than the firm. Other authors, e.g., Dickens et al. (1987), have taken the opposite approach, arguing that there are limits to bonding arrangements aside from the firm moral hazard. In particular, it is argued that incentive schemes which rely on disproportionately large punishments as compared to the crime they are designed to deter, cannot be implemented, i.e., such punishments will not hold up in a court of law. Still others, e.g., Malcolmson (1984), suggest that the problem can be resolved through the use of rank order tournaments, whereby the firm's compensation to its workforce is made state independent yet incentives for workers to put forth effort are preserved. We a priori rule out the possibility of utilizing third parties, to avoid the issue of the limitations placed on contracts through legal enforcement constraints. We also note that the option for the firm to default on its labor contract is typically ignored as a possibility in the tournament literature. Thus, our approach can be viewed as filling in this gap.

[5] See Arvan and Esfahani (1987).

[6] This assumption is also made in Arvan and Esfahani (1987). See note 8.

[7] In Arvan and Esfahani (1987) we show that there is no loss in generality to restricting attention to the case of full bond expropriation, since the maximum credible punishment will be applied to the firm when it expropriates, regardless of how many workers lose their bond.

expropriation has taken place.[8] Then, workers will be reluctant to post bond unless they believe that the value of the firm given that the firm honors the implicit contract exceeds the value of the firm given that the firm reneges via bond expropriation.

We question the viability of the reputation mechanism by considering a stochastic environment with information asymmetries. We assume that the firm experiences random productivity shocks. These productivity shocks affect the value of the firm. Workers are assumed to be unable to determine whether default is an attractive option to the firm, because workers cannot observe the realization of the current shock. In this respect our model borrows from the literature on implicit contracts under asymmetric information.[9] It turns out that in our model bond expropriation is an attractive option to the firm when it is actually in the low productivity state but claims to be in the high productivity state. The optimal contract must either convince workers that bond expropriation is not a threat, by giving a signal to workers about the firm's true state, or compensate them for the possibility of expropriation risk. When the former is the case, our model appears similar to the model of Milgrom and Roberts (1986), who analyze price and advertising signals of product quality in an experience good market. Their achievement is to consider a multidimensional signal, the product-price and advertising-expenditure pair, to explain the presence of noninformative advertising. Signaling in our model is also achieved via a multidimensional signal; the wage, performance bond, and employment triple. In addition to viewing contract offers as signals, our model explicitly considers both the worker moral hazard concerning shirking and the firm moral hazard concerning bond expropriation. A further feature which differentiates our approach from Milgrom and Roberts is that they only consider one shot uncertainty while we allow for an underlying stochastic process to generate workers' uncertainty. In Milgrom and Roberts' paper, all doubt about the firm's underlying characteristics is eliminated

8 When detection of bond expropriation is both instantaneous and certain, the firm going out of business constitutes the maximum credible punishment that workers can impose on the firm, as we showed in our previous paper. Weaker forms of punishment, such as increased recruiting costs, will sustain equilibria which are Pareto dominated by the equilibria described in this paper.

9 See the entire QJE (1983) supplement and Hart(1983) for a very good survey of this literature.

once buyers learn about the product quality associated with the firm in question. Thus, as they readily admit, their model is best interpreted as explaining noninformative advertising for a new product. In our model, even if workers are currently sure about the firm's productivity because the current contract offer is a perfectly informative signal, workers will be genuinely uncertain about the value of the firm in the next period because there is a new shock in that period. Indeed, our approach could be readily applied to the product quality issue to explain persistent advertising.

The basic model we consider is essentially an extension of the cogent efficiency wage model developed by Shapiro and Stiglitz (1984). We extend their model to a stochastic contract game which allows for the possibility of performance bonds, and hence for an analysis of the firm moral hazard problem. We view the problem of long term contract determination as one of constructing a sequential equilibrium for this contract game.[10] In fact, we restrict attention to the particular refinement of sequential equilibrium given by the *intuitive criterion*.[11] This solution concept is suitable since the issue of whether workers are willing to post performance bonds depends crucially on their beliefs about firm profitability. We assume that the underlying uncertainty of the game is generated by a very simple Markov process. This allows us to provide an explicit characterization of the equilibrium.

As a consequence of the workers' uncertainty about firm value, the firm does not necessarily rely exclusively on performance bonds to provide workers with the appropriate incentives when its productivity, and consequently its employment, is high. Instead, the firm may utilize efficiency wages, in part. The reliance on efficiency wages occurs when firm employment is sufficiently great that there would be a risk of bond expropriation, were the firm in the low productivity state. Performance bonds will still be utilized by the firm in such an equilibrium, but not to the extent that they would be were the firm's productivity observable. Moreover, the greater the value of the firm's shock in the high productivity state the greater the volume of

---

[10] See Kreps and Wilson (1982).

[11] See Cho and Kreps (1987).

employment in the high productivity state, and hence, the greater the reliance on efficiency wages. Thus, our model predicts a positive correlation between wages and firm size, a well known stylized fact of the labor market that heretofore has not been explained adequately.[12]

As an alternative to paying efficiency wages to signal high firm productivity, the firm might find it preferable to not signal its productivity at all, but instead to compensate workers for expropriation risk. This tactic is more favorable for the firm when the workers' prior beliefs put a lot of weight on the firm being in the high productivity state, since the size of the compensating differential is small in this case. This suggests that when workers are sufficiently optimistic, pooling rather than separating equilibria will prevail. However, as we show in the paper, this intuition is not correct. As long as worker prior beliefs are not held with certainty, the firm in the high productivity state has sufficient incentive to signal its productivity that any pooling equilibrium involving compensating differentials is undermined.

In addition to explaining why firms pay efficiency wages, our model resolves another paradox which has troubled the efficiency wage literature: When firms pay efficiency wages to workers, why don't they charge entry fees up front in order to recapture the rents embedded in the efficiency wage payments? More importantly, when such entry fees are utilized, doesn't the *involuntary* unemployment associated with efficiency wage theory vanish, since each worker's lifetime compensation, inclusive of the entry fee, equals what the worker could earn at his next best opportunity?[13] In our model, we allow the firm to charge each worker such an up front payment, as per Carmichael's suggestion. Yet in our signaling equilibrium, the marginal value product of the firm in the high productivity state exceeds the worker's reservation wage. In fact, this marginal value product even exceeds the efficiency wage! Thus, we demonstrate that the

---

[12] Standard efficiency wage theory argues that monitoring intensity varies inversely with firm size. For example see Katz (1986). This may indeed be the case. But such a negative correlation need not imply a positive correlation between firm size and wages, when one allows that performance bonds are viable. In addition, there does not appear to be a compelling reason for assuming that small firms have a comparative advantage in monitoring. Our explanation holds true under the assumption that monitoring intensity is constant, regardless of firm size, and remains true under the assumption that monitoring intensity varies inversely with firm size.

[13] See the exchange between Carmichael (1985) and Shapiro and Stiglitz (1985).

charging of up front fees is compatible with involuntary unemployment. Our result follows because the firm must convince its employees that there is no risk of bond expropriation. This places the following signaling constraint on the firm when in the high productivity state. The equilibrium payoff to the firm when in the low productivity state should be at least as large as the payoff that the firm in the low productivity state would obtain were it to offer the equilibrium contract for the firm in the high productivity state followed by it expropriating the performance bonds associated with this contract. This signaling constraint appears as an upward sloping curve in employment-wage space. That is, the firm in the high productivity state acts as if its supply of labor is not perfectly elastic, as long as the contract it offers calls for some degree of performance bonds. Furthermore, the firm in the high productivity state does not charge an entry fee to its employees, since doing so would only tighten the signaling constraint. These fees are charged to workers when the firm is in the low productivity state. The higher the efficiency wage paid in the high productivity state, the higher the fee collected in the low productivity state. However, sequential rationality imposes the restriction that the firm does not internalize this benefit. This is the crucial insight of the paper.

The rest of this paper is organized as follows. In section II we provide a simplified, one-period version of the model to provide some intuition for how efficiency wages work to signal firm value. In section III we provide the basic set up of the general, infinite horizon model. In section IV we discuss some of the main properties of separating equilibrium. In section V we characterize equilibrium. Finally, we offer a brief conclusion in section VI.

II. A Diagrammatic Approach

In this section we provide a simplified, one period version of our model to give an idea of how efficiency wages are used to signal firm value and to suggest those cases where such a signal will be utilized. Suppose a firm can be in one of two possible states, either m or M, where state m is less favorable to the firm than state M. Assume that the realization of the state is private

information held by the firm, i.e., workers cannot observe the firm's state. Let U_μ denote the lifetime value of the firm when it is in state μ for $\mu = m, M$. $U_m < U_M$. Assume the firm's only input is labor and that the firm's choice problem is to select the profit maximizing employment contract from among those contracts which workers will find acceptable. In this simplified version of our model, a contract is a pair, (W, L), where W is the contract wage and L is the volume of employment.

Workers can earn the reservation wage, W_R, by seeking employment elsewhere. Hence, $W \geq W_R$ is a necessary condition for contract acceptability. In addition, suppose that employment with the firm necessitates that workers post a performance bond of value B. The other necessary condition which governs contract acceptability is that workers must believe that the firm won't expropriate the performance bonds that they post. We elaborate on this condition below.

Suppose that if the contract (W, L) is accepted then the current period profit of the firm in state μ is $\Pi(\mu, W, L)$ and suppose that if, in addition, the firm expropriates the performance bonds then its lifetime value is $\Pi(\mu, W, L) + BL$, for $\mu = m, M$. If workers were primarily concerned about expropriation by the firm when its state is m, they would not accept a contract which satisfied: $U_m < \Pi(m, W, L) + BL$. When this inequality is reversed workers should be convinced that there is no risk of bond expropriation. We refer to the constraint: $U_m \geq \Pi(m, W, L)$ as the signaling constraint, SC, because when it is violated workers beliefs will assign high probability to the firm being in state m. The firm can only convince its workers that it is in state M by offering a contract which satisfies the SC.

Assume that Π exhibits the following properties: $\Pi_\mu > 0$, $\Pi_W < 0$, $\Pi_{LL} < 0$, and $\Pi_{\mu L} > 0$. Let $L_M^* = \underset{L}{\operatorname{argmax}}\ \Pi(M, W_R, L)$. We focus on the problem which determines the firm's optimal contract when the firm is in state M, since in equilibrium it is only in this state that efficiency wages are possible. This problem is given by

$$\underset{W \geq W_R,\ L \geq 0}{\text{maximize}} \quad \Pi(M, W, L) \qquad \text{subject to} \quad U_m \geq \Pi(m, W, L) + BL$$

In figure 1 we graph the solution to this problem under the assumption that the SC does not bind. Observe that the isoprofit curve is tangent to the horizontal line $W = W_R$ at $\overset{*}{L}_M$ and the SC cuts the horizontal line $W = W_R$ at $L_0 > \overset{*}{L}_M$. This is necessary when the SC doesn't bind.

In figures 2 and 3 we graph the solution to this problem when the SC binds. We identify two candidates as solutions of this problem. The appropriate choice depends on the respective slopes of the isoprofit curves for the firm in state M and the SC. In figure 2, there is an interior solution. This is the case where efficiency wages occur. The first order necessary condition in this case is $\Pi_L(M, W_M, L_M) = \Pi_L(m, W_M, L_M) + B$. The right hand side of this equation is positive when the SC is rising. When Π_L is interpreted as the difference between the marginal value product of labor and the wage and when the tangency point occurs on the rising portion of the SC, the marginal value product of labor exceeds the efficiency wage, W_M! In this sense our model predicts even grosser allocative inefficiency than the standard efficiency wage model, such as Shapiro and Stiglitz (1984). In figure 3, there is a corner solution, $(W_M, L_M) = (W_R, L_0)$. Efficiency wages are not utilized in this case, though there is evidently allocative inefficiency relative to the first best solution where workers can observe the firm's state. The first order necessary condition in this case, $\Pi_L(M, W_R, L_0) \leq \Pi_L(m, W_R, L_0) + B$, indicates that the firm finds it too costly to utilize efficiency wages to expand its employment.

In the subsequent sections we expand on this basic model in order to endogenize those values which were taken parametrically in this simplified version. However, the main message remains intact when going to the full model: Efficiency wages are utilized to signal the firm's state and thereby to convince workers that there is no risk of bond expropriation. This signaling distorts the labor allocation relative to the first best solution. Moreover, we are able to show that this distortion remains, even when the firm charges its employees up front fees, so that over the lifetime of employment workers only earn their reservation wage.

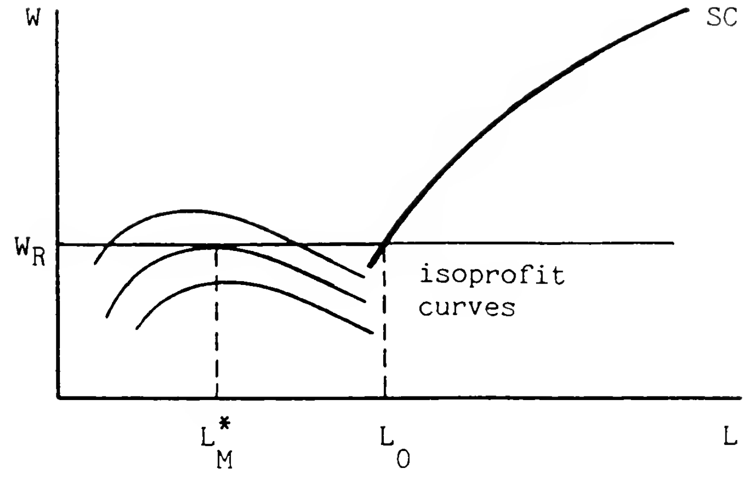

Figure 1

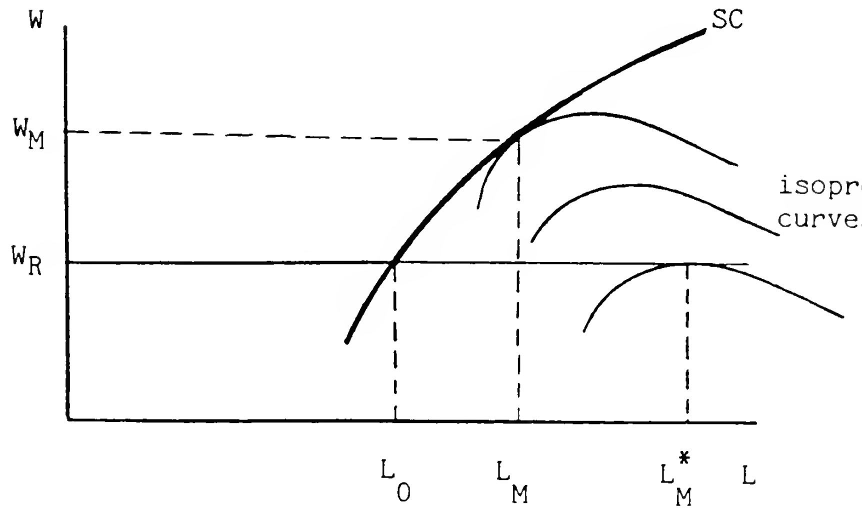

Figure 2

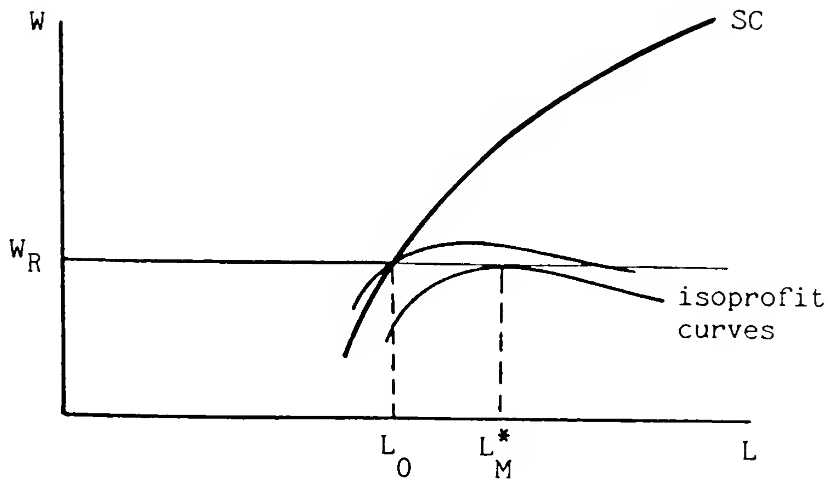

Figure 3

III. Preliminaries

There are three commodities in the model: a produced good, labor-leisure, and a stock commodity which serves as numeraire. There are two types of agents, the firm and workers. Workers are identical in every respect except employment status. Each worker is infinitely lived and endowed with both an indivisible unit of labor-leisure and enough of the numeraire stock to be able to post the requisite bond, should the worker choose to do so. Worker preferences are specified by the rate of time preference, r, which coincides with the market interest rate, and a per period von Neumann-Morgenstern utility function defined over leisure and numeraire consumption. As is typical in this literature, we assume that workers are risk neutral. Workers who are employed by the firm get to choose their level of effort, $e \in [0,1]$. When $e = 0$ an employed worker is taking leisure on the job or shirking. When $e = 1$ the worker is being productive or putting forth effort. Intermediate values of e are interpreted as either the fraction of the time at work which the worker puts forth effort or the probability that the worker does not shirk. The utility function is normalized so that when the worker is consuming y units of numeraire and taking e units of effort, the per period utility is $y-e$. The time invariant, discounted, lifetime expected utility at the worker's best alternative employment is denoted by V_u and is assumed to be a nonnegative number.

The production function of the firm, f, is defined over the effective labor input L^e. In what follows all workers will be choosing the same effort level, in which case $L^e = eL$, where e is the common effort level. It is assumed that f is increasing, differentiable, and strictly concave. In addition, $f(0)=0$, $\lim_{L \downarrow 0} f'(L) = \infty$, and $\lim_{L \uparrow \infty} f'(L) = 0$. The firm is subject to a random, privately observed, internal price. Let μ_t be the random variable denoting this price in period t. We assume that μ_t takes on only two values, either m or M, with $0 < m < M$. We also assume that the stochastic process $\{\mu_t\}_{t=0}^{t=\infty}$ is a Markov chain with transition probabilities given by

$Pr\{u_{t+1} = m | u_t = M\} = a$ and $Pr\{u_{t+1} = m | u_t = m\} = b$. Assume $0 < a < b < 1$. The firm is taken to maximize discounted expected profits, where the discount rate is the market interest rate, r.

Each period that the firm is in operation is broken down into four stages of decision. In the first stage, the firm offers a contract. A contract in period t is an ordered 5-tuple, $\boldsymbol{C}_t$, $\boldsymbol{C}_t = (W_t, B_t, L_t, F_t, R_t)$, where W_t is the wage rate to be paid to each employed worker, B_t is the performance bond to be posted by each employed worker, L_t is the volume of employment should the contract be accepted, F_t is the fee paid by workers who are in the firm's reserve pool, and R_t is the size of the reserve pool. The pool of workers to whom the firm may offer a contract in period t, in order of priority, consists of: first, those workers employed by the firm in period $t-1$ who are in good standing; second, those workers in the firm's reserve pool in period $t-1$; and third, the outside workers who are unemployed in period t and seek employment from the firm in that period. Within each of these priority classification jobs are allocated on a pro rata basis.

In the second stage each worker who is offered the contract either accepts it or rejects it. The contract is null and void in the event that the firm cannot find L_t workers to accept it. In this case, the remaining two stages are foregone and the firm necessarily continues operation into period $t+1$. Any worker who did accept the contract is freed from this obligation and entitled to seek employment elsewhere in period t. Similarly, each worker who is offered a slot in the reserve pool either accepts it or rejects it. Acceptance requires payment of the fee, F_t, which is nonrefundable. Workers in the reserve pool can obtain employment elsewhere in period t, i.e., they can earn $rV_u/(1-r)$ in the period. Indeed, our reason for including a reserve pool in the model is that when labor contracts call for efficiency wages, workers earn rents and thus should be willing to pay for the right to employment.[14] The firm could extract such a payment by hoarding workers, encouraging them to shirk, and capturing the performance bonds of those workers it caught shirking. When the workers' reservation wage is positive, this is an

[14] See Carmichael (1985).

inefficient mechanism for extracting such payments. The reserve pool mechanism, where workers pay for the right to be offered employment when vacancies appear but may work elsewhere in the meantime, is superior.

The contract is operational when L_t workers accept it. Once the contract is operational, each of the L_t employees receives the wage W_t and posts the bond B_t. Thus the net payment made by the firm to each employee at this juncture is $W_t - B_t$, which may be positive or negative. In the third stage, each worker chooses e_t. This choice is not observed by the firm. Instead, the firm has a chance of detecting the worker taking leisure on the job. The detection probability is $(1-e_t)\theta$, where θ is a parameter reflecting the firm's monitoring intensity, $0 < \theta \leq 1$. In the concluding section we briefly consider the firm's problem when it is free to choose the magnitude of θ.

Once the third stage has occurred, production occurs and output is publicly observed. Both the firm and its workers are able to impute the effective labor input in period t, L_t^e, via observation of firm output and their knowledge of the production function, f. We assume that there is a deterministic relationship between L_t, L_t^e, and, the number of workers detected shirking in period t, L_t^d. This relationship is given by $L_t^d = (L_t - L_t^e)\theta$. The reason for making this assumption is that when there are layoffs, $L_{t+1} < L_t - L_t^d$, the retention probability is nonlinear in L_t^d. Hence, the value of future employment to workers currently under contract could depend on noise in the detection of worker shirking, i.e., if L_t^d were a random variable then the value of future employment would depend on the distribution of this random variable. This effect could conceivably offset our argument against labor hoarding.

Given L_t^d, the firm decides in the fourth stage to either continue operation into period $t+1$ or cease operation at the end of period t. When the firm has decided to continue operation, all the workers who are detected shirking are fired and made to forfeit their performance bonds to the firm. All of the other $L_t - L_t^d$ workers are considered employees in good standing. At the start

of period $t+1$ such employees get back $(1+r)B_t$ from the firm. When the firm decides to cease operation at the end of period t, observation of L_t^d is immaterial. In this case, the firm keeps the performance bonds of all its workers, regardless of whether they have been detected shirking. The game ends once the firm has decided to cease operation.

Our goal is to construct sequential equilibria in strategies where the history of play can be completely summarized by workers' current beliefs about the firm's internal price. This restriction on strategies appears natural given the Markovian structure of the underlying uncertainty.[15] In order to construct such equilibria, we must first specify what strategies and beliefs are and then impose sequential rationality, on strategies, and consistency with Bayes' Law and with equilibrium strategies, on beliefs. We shall do so in a rather informal manner, to save on notation and thereby to enhance readability.

At the start of period t, workers will have genuine uncertainty as to the current value of the firm's internal price. Let α_t represent the beliefs of workers that the firm is in the low productivity state in period t, before the contract in period t is offered but after the play of the game through period $t-1$ has been completed: $\alpha_t = \Pr\{\mu_t = m\}$. The contract offer serves as a signal of the firm's internal price. Let β_t represent these revised beliefs. Then $\beta_t = \beta(\alpha_t, C_t)$, i.e., the initial beliefs and contract offer map into revised beliefs. The consistency requirement restricts β via Bayes' Law for the equilibrium values of C_t, but imposes no restriction on β for out of equilibrium values of C_t. If the contract offer is not accepted by L_t workers, then Bayes' Law requires that $\alpha_{t+1} = \alpha_t b + (1-\alpha_t)a$. If the contract in period t is operational, then continuation of firm operation into period $t+1$ may also signal the firm's internal price in period t and hence provide information that is relevant in predicting the firm's internal price in period $t+1$. Then Bayes' Law requires:

$$\alpha_{t+1} = \frac{\beta_t H_{tm}\, b + (1-\beta_t) H_{tM}\, a}{\beta_t H_{tm} + (1-\beta_t) H_{tM}}.$$

---

[15] Note, however, that this restriction on strategies is not innocuous. The restriction has the effect of taking away all the workers' bargaining power. In particular, under the restriction workers can't play trigger strategies where they punish the firm, by not accepting the contract offer, unless the contract provides employees with sufficient rents.

where $H_{t\mu}$ denotes the probability that the firm honors the contract given $\{\mu_t = \mu\}$, for $\mu = m, M$.

A strategy for a worker is an ordered triple of functions. The first function maps initial beliefs and contract offers into acceptance or rejection decisions. This function governs the employment decision. The second function also maps initial beliefs and contract offers into acceptances or rejections, but this one governs the decision concerning entry into the reserve pool. The third function maps revised beliefs and contract offers into effort levels. Thus the first two functions govern play in stage two, while the third function governs play in stage three. In what follows we assume that all workers pursue the same strategy.

A strategy for the firm is an ordered pair of functions. The first function maps the firm's internal price and workers' initial beliefs into contract offers. The second function maps the firm's internal price, workers' revised beliefs, and contract offers into the probability of honoring the contract. The first function governs play in stage one, while the second function governs play in stage four.

IV. Separating Equilibrium

Suppose equilibrium of the game is characterized by contract offers which are independent of worker initial beliefs, at least for $\alpha_t \in (0,1]$. Letting C^* denote the equilibrium contract offer function, we have $C^*(m,\alpha_t) = C_m$ and $C^*(M,\alpha_t) = C_M$ for all $\alpha_t > 0$. Moreover, suppose the equilibrium is separating, i.e., $C_m \neq C_M$. Then, $\beta(\alpha_t, C_m) = 1$ and $\beta(\alpha_t, C_M) = 0$ for $\alpha_t > 0$. That is, upon observation of the equilibrium contract offers, workers act as if they know the realization of the firm's internal price. Also assume that the equilibrium contract offers are always accepted.

Before continuing further, we introduce some notation to aid the exposition. Let $\Pi(\mu, C_t) = \mu f(L_t) - W_t L_t + F_t R_t$. The function Π gives current period profit for a firm with internal price μ when it offers contract C_t, which has been accepted and at which employees put forth

effort. Let $\pi_m = \Pi(m, C_m)$ and $\pi_M = \Pi(M, C_M)$. Thus, π_m and π_M denote equilibrium current period profits for the firm with internal price equal to m and M, respectively. That is, we assume employees do put forth effort in equilibrium.[16] Let V_m and V_M denote the equilibrium, lifetime, discounted expected values of the firm with current internal price equal to m and M, respectively. These values are given by

$$(1) \qquad V_m = \frac{(1+r)[(r+a)\pi_m + (1-b)\pi_M]}{r\,[1+r+a-b]} \quad \text{and}$$

$$V_M = \frac{(1+r)[a\pi_m + (1+r-b)\pi_M]}{r\,[1+r+a-b]},$$

under the assumption that the firm always honors the contract in equilibrium.[17]

If the firm with internal price equal to m offered an unacceptable contract, it would earn zero profits in the current period and its lifetime expected value would be $[b V_m + (1-b) V_M]/(1+r)$. Since $V_m = \pi_m + [b V_m + (1-b) V_M]/(1+r)$, individual rationality requires that $\pi_m \geq 0$. Similarly, individual rationality requires $\pi_M \geq 0$. Thus, either $\pi_m = \pi_M = 0$ and consequently $V_m = V_M = 0$, in which case the equilibrium is trivial with the firm never in operation, or $\max\{\pi_m, \pi_M\} > 0$, in which case $V_m, V_M > 0$. In fact, since $\Pi(M, C_m) \geq \Pi(m, C_m)$, with strict inequality when $L_m > 0$, and since $\pi_M \geq \Pi(M, C_m)$, because C_M is a best response for the firm when its internal price is M, it must be that $\max\{\pi_m, \pi_M\} = \pi_M$. Hence, in a nontrivial equilibrium, $V_M > V_m > 0$. This is the case we focus on below.

Consider the firm's stage four decision given that an arbitrary contract has been offered and accepted in period t and that the internal price is m. Were it to cease operation at the end of period t, the firm would expropriate the performance bonds of all its workers who would otherwise be in good standing. The value of this expropriated bond is $B_t(L_t - L_t^d) =$

[16] This is shown to be a necessary property of equilibrium in lemma 1.

[17] This is also shown to be a necessary property of equilibrium in lemma 1.

$B_t L_t[1-(1-e_t)\theta]$. Were the firm to continue operation, the firm's lifetime, expected, discounted value equals $[b\,U_m + (1-b)\,U_M]/(1+r)$. Thus, the firm honors the contract only if

$$(2) \quad b\,U_m + (1-b)\,U_M \geq (1+r)B_t L_t[1-(1-e_t)\theta].$$

(2) is termed the no expropriation condition for the firm in state m, NEC$_m$, where the subscript refers to the current internal price of the firm. NEC$_M$ is defined similarly for the firm with current internal price equal to M. Since $a < b$, the NEC$_m$ is more restrictive than the NEC$_M$. The model was constructed so that this would be the case. Note that the expropriation decision is independent of W_t, F_t, and R_t, since these variables are "sunk" by the third stage and have no impact on either the effort decision or on the future profits of the firm. For a contract such that the NEC$_m$ is violated but the NEC$_M$ is satisfied, worker beliefs are critical in determining the extent of expropriation risk, and hence, the value of accepting the contract $\boldsymbol{C}_t$. This may provide a motive for the firm to signal its workers when its internal price is M. It is this motive which explains why the firm might be willing to pay efficiency wages.

From the workers' point of view, the effort decision in stage three depends on the current disutility of effort, worker beliefs about the productivity of the firm, and the expected gains from being in good standing at the end of period t. Let the equilibrium lifetime expected utility of a worker employed with the firm be V_m and V_M, when the current internal price of the firm is m and M, respectively. Also, let the equilibrium employment levels of the firm be L_m and L_M in states m and M, respectively. Then the lifetime, discounted expected value of an employee under the contract $\boldsymbol{C}_t$ who has chosen the effort level e_t is given by

$$(3)\; V(e_t, \boldsymbol{C}_t) = W_t - e_t + \{\theta(1-e_t) + [1-\theta(1-e_t)]\,p_{texp}\}[\frac{V_u}{1+r} - B_t] + [1-\theta(1-e_t)][1-p_{texp}]$$

$$\times \; \frac{\alpha_{t+1}[s_{tm}\,V_m + (1-s_{tm})\,V_u] + (1-\alpha_{t+1})[s_{tM}\,V_M + (1-s_{tM})\,V_u]}{1+r},$$

where p_{texp} is the probability that the firm reneges on the contract via expropriation, $\Lambda_{texp} = \beta_t(1-H_{tm}) + (1-\beta_t)(1-H_{tM})$, s_{tm} is the probability that a worker in good standing after period t is retained by the firm given $\{u_{t+1}=m\}$, $s_{tm} = \min\{L_m/[1-\theta(1-e_t)]L_t, 1\}$, and s_{tM} is defined similarly given $\{u_{t+1}=M\}$. Note that the individual employee takes s_{tm} and s_{tM} as parameters. The effort level used to evaluate these parameters is the common effort of the employee's coworkers. To better understand (3), consider the expression term by term. The first term is the wage minus the disutility of effort. The second term is the product of the probability of losing the performance bond, both through detection of shirking and through expropriation by the firm, and the difference between the value of being employed elsewhere in the subsequent period and the value of the lost bond. The third term is the product of the probability that the worker is in good standing in period $t+1$ given that the firm operates in that period, the probability that the firm operates in period $t+1$, and the expected value to a worker who is in good standing when the firm operates in period $t+1$.

Each employee under contract chooses e_t to maximize (3). Observe that (3) is linear in e_t so that there is always a corner optimum. The condition that $e_t=1$ yields a maximum is termed the No Shirking Condition, NSC, and is given by

$$(4)\quad -p_{texp}B_t + (1-p_{texp})\frac{\alpha_{t+1}s_{tm}(V_m - V_u) + \alpha_{t+1}s_{tM}(V_M - V_u)}{1+r} \geq \frac{1}{\theta} - B_t.$$

The first component of a worker's strategy is governed by the employment acceptance condition, EAC. This condition is given by

$$(5)\qquad \max[V(0,C_t), V(1,C_t)] \geq V_u.$$

Workers will not accept an employment offer when (5) is violated.

The second component of a worker's strategy is governed by the reserve pool acceptance condition, RAC. Since there is no expropriation risk for those workers who are offered a position in the firm's reserve pool, this condition is given by

$$(6)\quad (1-p_{texp})\,\frac{\alpha_{t+1}g_{tm}(V_m - V_u) + (1-\alpha_{t+1})\,g_{tM}(V_M - V_u)}{1+r} \geq F_t,$$

where $g_{t\mu}$ is the probability that a worker who is in the reserve pool in period t gains employment from the firm which is in state μ in period $t+1$, for $\mu = m, M$: $g_{t\mu} = \min\{1, \max[0, (L_\mu - L_t + L_t^d)/R_t]\}$

Let e_m and e_M denote the equilibrium effort levels and let H_m and H_M denote the equilibrium probabilities that the firm honors the contract in states m and M, respectively. As we have already noted, $e_m = e_M = H_m = H_M = 1$. Below we demonstrate why this is necessarily the case. It is simple to see that $H_m, H_M \neq 0$. For were this not the case, then from (6) it is apparent that no worker would be willing to pay to enter the reserve pool. Moreover, from (4) it is just as apparent that all employed workers would shirk. Hence, the EAC reduces to $W_t - B_t \geq rV_u/(1+r) \geq 0$ in this case. But since firm profit per employee would be $B_t - W_t$ and firm operation would cease after period t, the firm would be better off offering an unacceptable contract in period t so that it could continue operation into period $t+1$.

In lemma 1 we show that intermediate values of H_m and H_M can be ruled out as well. The intuition for this result is that were such randomizing optimal, the firm would necessarily be indifferent between continuing operation and expropriating the bonds. But when the firm randomizes over its decision to expropriate the performance bonds or honor the contract, its employees face some expropriation risk. As a result, a compensating differential must be embedded in the contract wage, to induce the employees to accept the expropriation risk. The firm could offer an alternate contract, by reducing the size of the bonds posted by its workers, thereby eliminating the expropriation risk and, consequently, lowering the wage which workers find

acceptable. When the alternate contract has the property that it is less profitable to the firm, were it in the other state, than the equilibrium contract offered by the firm when in the other state, workers' revised beliefs should be the same whether the original contract or the alternate contract is offered. This is precisely the requirement that the intuitive criterion imposes.[18] We demonstrate that an alternate contract with this property exists. Therefore, by invoking the intuitive criterion, an equilibrium with partial bond expropriation and an associated compensating differential would be undermined through the offering of such an alternate contract.

Since the firm honors the equilibrium contract and since the NEC_m is more restrictive than the NEC_M, workers will not perceive expropriation risk at the contract $\boldsymbol{C}_m$, regardless of their beliefs. If $V_m > V_u$, then the firm should be able to lower W_m while keeping the contract in state m acceptable, as long as workers' revised beliefs were unaffected. We show that this can be done, again by invoking the intuitive criterion. Therefore, an equilibrium with $V_m > V_u$ would be undermined. Note that this argument is not applicable to rule out the possibility that $V_M > V_u$, when workers would perceive expropriation risk at the contract $\boldsymbol{C}_M$ were it offered by the firm when its internal price equals m.

<u>Lemma 1</u>. Suppose the separating equilibrium satisfies the intuitive criterion. Then (i) $H_m = H_M = 1$, (ii) $V_m = V_u$, and (iii) if $V_u > 0$, then $e_m = e_M = 1$.

<u>Proof</u>: We first show that $H_m = 1$ and that if $V_u > 0$, then $e_m = 1$. Suppose not. Consider an alternate contract constructed from $\boldsymbol{C}_m$, call it $\hat{\boldsymbol{C}}_m$, where $\hat{W}_m = W_m$, $\hat{B}_m = H_m B_m + \varepsilon$, for some $\varepsilon > 0$, $\hat{L}_m = e_m L_m$, $\hat{F}_m = [F_m R_m + (\theta B_m - W_m)(1-e_m)L_m]/[R_m + (1-e_m)L_m]$, and $\hat{R}_m = R_m + (1-e_m)L_m$. In essence, the contract $\hat{\boldsymbol{C}}_m$ has a slightly larger effective bond than $H_m B_m$, the effective bond resulting from $\boldsymbol{C}_m$, without the expropriation risk. It also places

[18] Cho and Kreps explain the intuitive criterion as follows:

> I am sending message m' which ought to convince you that I know t'. For I would never wish to send m' if I know t, while if I know t', and if sending this message so convinces you, then as you can see, it is in my interest to send it.

the workers who shirk under C_m into the reserve pool, along with those workers already in the reserve pool under C_m, and then averages the fee the firm is collecting from the former, $\theta B_m - W_m$, with F_m. Note that when $\hat{C}_m$ is acceptable and when workers employed under $\hat{C}_m$ put forth effort, the effective labor input and the wage rate under both contracts are the same. In this case the firm has the same lifetime discounted value when it offers C_m as when it offers $\hat{C}_m$, regardless of its internal price.

Since either $H_m < 1$ or $e_m < 1$, $\hat{B}_m \hat{L}_m < [bU_m + (1-b)U_M]/(1+r)$ for ε sufficiently small, i.e., there is no expropriation risk under $\hat{C}_m$. Moreover, $\hat{s}_\mu \geq \min\{L_\mu / e_m L_m, 1\} \geq \min\{L_\mu/[1-\theta(1-e_m)]L_m, 1\} = s_\mu$, for $\mu = m, M$, since $e_m \leq 1-\theta(1-e_m)$. Thus, when $e_m > 0$ and $\beta(\alpha_i, \hat{C}_m) = 1$, the NSC is satisfied as a strict inequality under $\hat{C}_m$ and, therefore, employees necessarily put forth effort under this contract.

Since C_M is optimal when the firm's internal price is M, it must be that

$$\text{(7)} \qquad \pi_M \geq M f(e_m L_m) - W_m L_m + F_m R_m .$$

When (7) holds as a strict inequality, which we assume for now, the intuitive criterion requires that $\beta(\alpha_i, \hat{C}_m) = 1$. The case where (7) holds as an equality is considered later in the proof.

When $e_m > 0$, the EAC for the contract C_m can be written as

$$\text{(8)} \qquad W_m - 1 - (1-H_m)B_m + \frac{bH_m(V_m - V_u)}{1+r} + \frac{(1-b)H_m(V_M - V_u)}{1+r}\min\{1, \frac{L_M}{[1-(1-e_m)\theta]L_m}\} \geq \frac{rV_u}{1+r} .$$

When $e_m < 1$, the EAC for the contract C_m can be written as

$$(9) \qquad -(1-\theta)(1-H_m)B_m + \frac{(1-\theta)bH_m(V_m - V_u)}{1+r} +$$

$$\frac{(1-\theta)(1-b)H_m(V_M - V_u)}{1+r}\min\{1, \frac{L_M}{[1-(1-e_m)\theta]L_m}\} \geq \frac{rV_u}{1+r} + \theta B_m - W_m$$

Similarly, the RAC for the contract $\boldsymbol{C_m}$ can be written as

$$(10) \qquad \frac{bH_m(1-e_m)L_m(V_m - V_u)}{R_m(1+r)} +$$

$$\frac{(1-b)H_m \min\{R_m, \max[0, L_M - [1-(1-e_m)\theta]L_m]\}(V_M - V_u)}{R_m(1+r)} \geq F_m$$

We will now show that (8)-(10) imply contract acceptability for $\hat{\boldsymbol{C}}_m$. Since there is no shirking and no expropriation under $\hat{\boldsymbol{C}}_m$ and since $[1-(1-e_m)\theta]L_m \geq e_mL_m$, (8) implies the EAC for $\hat{\boldsymbol{C}}_m$. Indeed, this condition holds as a strict inequality when $H_m < 1$. Similarly, multiplying (9) by $(1-e_m)L_m/[R_m+(1-e_m)L_m]$, (10) by $R_m/[R_m+(1-e_m)L_m]$, and adding the results implies the RAC for $\hat{\boldsymbol{C}}_m$. In this case the condition holds as a strict inequality when $e_m < 1$ and either $H_m < 1$ or $V_u > 0$. When the EAC holds as a strict inequality the firm can offer an alternative contract which yields higher profits by offering a slightly lower wage than W_m. When the RAC holds as a strict inequality the firm can offer an alternative contract which yields higher profits by charging a slightly higher fee than F_m. Moreover, from the intuitive criterion workers should be certain that the internal price of the firm is m when they observe this alternate contract, since (7) holds as a strict inequality. This shows that $H_m = 1$ and $e_m = 1$ when $V_u > 0$. Similar reasoning rules out $V_m > V_u$, since were this the case (8) would hold as a strict inequality and the firm would have an incentive to upset the equilibrium by lowering its wage offer.

We turn to the case where (7) holds as an equality. Alter the contract $\hat{C}_m$ to the contract $\tilde{C}_m$, were the two contracts differ only in their wage rates and employment levels. $\tilde{L}_m = \hat{L}_m - dL$, where dL is small but positive, and $\tilde{W}_m = \hat{W}_m + [(m+M)f'(\hat{L}_m)/2 - \hat{W}_m]dL/\tilde{L}_m$. Then $\Pi(m, \tilde{C}_m) > \Pi(m, \hat{C}_m)$ and $\Pi(M, \tilde{C}_m) < \Pi(M, \hat{C}_m)$. Thus, the same argument as given above can be applied to the contract $\tilde{C}_m$.

A similar construction can be given to generate the contract $\hat{C}_M$. Observe that if $H_M < 1$, then $B_M L_M[1-\theta(1-e_m)] = [aU_m + (1-a)U_M]/(1+r) > [bU_m + (1-b)U_M]/(1+r)$. Thus, when $H_M < 1$ the firm prefers C_M to $\hat{C}_M$, given that its internal price is m, since $\hat{C}_M$ involves less bonds to expropriate than C_M. When $e_M < 1$, either the firm prefers C_M to $\hat{C}_M$, given that its internal price is m, again because $\hat{C}_M$ involves less bonds to expropriate than C_M, or an alternate contract, $\tilde{C}_M$, can be constructed so that $\Pi(m, \tilde{C}_M) < \Pi(m, \hat{C}_M)$ and $\Pi(M, \tilde{C}_M) > \Pi(M, \hat{C}_M)$. The construction is the same as the one given in the paragraph above except that in this case $\hat{L}_M < \tilde{L}_M$ is required. SS

Let W_R denote the workers' reservation wage inclusive of the disutility of effort, $W_R = 1 + rV_U/(1+r)$. Since $V_U = V_m \leq V_M$, it must be that $W_m \leq W_R \leq W_M$. Indeed, by taking account of the properties of equilibrium given in lemma 1, one can solve for the rent per worker resulting from being employed when the firm's internal price is M

$$(11) \quad V_M - V_U = \frac{(W_M - W_R)(1+r)}{r+a}$$

Let L_m^* be implicitly defined by $mf'(L_m^*) = W_R$. Note that $L_m \leq L_m^*$, even when $W_m < W_R$.[19] It follows that $L_m \leq L_M$; for if not, $\pi_M < \Pi(M, C_m)$. From (11) and the fact that (1-

---

[19] Were $L_m > L_m^*$, the firm could do better by lowering its employment to L_m^*, putting $L_m - L_m^*$ workers into the reserve pool when $L_M > L_m^*$, or cutting its wage by $(1-b)L_M(V_M - V_U)/(1+r)[1/L_m^* - 1/L_m]$

b) is the transition probability from state m to state M, the expected discounted rent per worker when the firm's internal price is m and the firm's employment is L is $\min\{L_M/L,1\}(1-b)(W_M-W_R)/(1+r)$.[20] By invoking an argument similar to the one given in the proof of lemma 1, there is no loss in assuming that the firm extracts all future rents accruing to its workers when its internal price is m. Then from (11), lemma 1, and assuming there are no redundant workers in the reserve pool we have

$$(12)\qquad F_m(L) = \min\{\frac{L_M}{L},1\}\frac{(1-b)(W_M-W_R)}{r+b}, \qquad W_m(L) = W_R - F_m(L),$$

$$B_m \le \frac{1}{\theta} - F_m(L), \text{ and } \qquad R_m(L) = \max\{L_M - L, 0\}.$$

Note that the weak inequality restricting B_m necessarily holds as and equality only when the NEC$_m$ binds. In essence, all that is left for the firm to choose when its internal price is m is its level of employment. Thus the problem which determines the firm's optimal contract when its internal price equals m can be written as

$$(13)\qquad \underset{L\ \ge 0}{\text{maximize}}\quad m f(L) - W_R L + F_m L_M \quad \text{Subject to } [\frac{1}{\theta} - F_m(L)]L \le \frac{bU_m + (1-b)U_M}{(1+r)}.$$

Since L_m is the solution to (13), it is evident that either $L_m = L_m^*$ or the NEC$_m$ binds in equilibrium. Note that in (13) $F_m L_M$ is the value of the rents which the firm can extract from its workers, both from its employees, in the form of wage reductions, and from the workers in its reserve pool, directly in the form of entry fees. Observe that this amount is fixed, regardless of the choice of employment level.

---

when $L_M \le L_m^*$. That such an alternative contract would be acceptable follows by a similar argument to the one given in the proof of lemma 1.

[20] It is appropriate to compute these rents from the point of view of play in stage three, after acceptance of the current contract has already occurred.

There is no reason for the firm to take any costly action in order to convince its work force that its internal price is m, both because workers don't earn any rents and because workers view bond expropriation to be more likely in this state. When the firm's internal price is M, the situation is reversed. There may very well be a reason for the firm to pay a cost in order to convince its employees of its true internal price. For workers to be so convinced, the wage-employment pair, (W, L), must satisfy the following signaling constraint, SC.

$$(14) \quad U_m \geq m f(L) - WL + \max\left\{\frac{b\,U_m + (1-b)\,U_M}{1+r}, B_M(L)L\right\},$$

where $B_M(L) = 1/\theta - F_M(L)$ and $F_M(L) = \min\{L_M/L, 1\}(1-a)(W_M - W_R)/(r+a)$, the expected discounted rent per worker when the firm's internal price is M and its employment is L. When (14) is satisfied, the firm has no incentive to offer the wage-employment pair (W, L) when in state m. Thus, when the contract satisfies the SC and the firm in the high productivity state benefits from offering such a contract, workers' revised beliefs should place all the weight on the firm being in the high productivity state, by the intuitive criterion. Observe that when $W_M > W_R$, that is, the firm utilizes efficiency wages, then $B_M < B_m$, i.e., there is less reliance on performance bonds when the firm's internal price is M than when the firm's internal price is m, as one's intuition would suggest. Note however, that $B_m < 1/\theta$ in this case.

The problem which determines the firm's optimal contract when its internal price is M is given by

$$(15) \quad \underset{L \geq 0,\, W \geq W_R - F_M(L)}{\text{maximize}} \quad M f(L) - WL$$

$$\text{Subject to: (14), the SC, and} \quad B_M(L)L \leq \frac{a\,U_m + (1-a)\,U_M}{(1+r)}$$

Note that the lower bound constraint on W in (15) is just the EAC. Observe that when the SC binds, the lower bound constraint on W may not bind. This is the case where efficiency wages are

utilized. It seems reasonable to view the signaling constraint, where it might be binding, as an upward sloping curve in employment-wage space, at least when the contract violates the NEC_m.[21] The slope of an isoprofit curve for the firm with current internal price equal to M is given by

$$(16)\qquad \frac{dW}{dL} = \frac{M f'(L) - W}{L}.$$

The slope of the SC is given by

$$(17)\qquad \frac{dW}{dL} = \frac{m f'(L) - W + B_M}{L}.\ ^{22}$$

In order to restrict attention to the case where this intuition applies, we make the following assumption.

<u>Assumption 1.</u> $\left[\frac{1+r}{r+a} - \frac{m}{M} \right] W_M < \frac{1}{\theta} - \frac{1-a}{r+a} W_R$

That is, we restrict attention to equilibrium values of W_M which satisfy an upper bound constraint. It is straightforward to show that under assumption 1, $(W_M - B_M)/m < W_M/M$. Thus, it follows from (16) and (17) that when the equilibrium is given by the tangency of an isoprofit curve and the SC, both curves must be rising, as long as assumption 1 is satisfied. This assumption is innocuous for low values of θ but is restrictive when θ is near 1.

V. Characterizing the Equilibrium

<u>IV.1 Full Bonding Equilibrium</u>

It is evident from (15) that when the SC does not bind in equilibrium then the CAC must bind, i.e., $W_M = W_R - F_M$, by invoking the same argument as the one used in the proof of lemma

---

[21] Note that when we graph the SC we treat W rather than L as the dependent variable.

[22] The slope of the SC given in the text is only correct for $L < L_M$. When $L > L_M$ then along the SC we have $\frac{dW}{dL} = \frac{m f(L) - W + 1/\theta}{L}$

1. When this occurs it must be that $F_M = 0$ and $W_M = W_R$. In this case $B_m, B_M = 1/\theta$. We refer to such an equilibrium as a full bonding equilibrium, since all of the employees' incentive to put forth effort is provided via performance bonds. In a full bonding equilibrium, $V_m = V_M = V_U$. Thus, one would not associate such a contractual equilibrium with involuntary unemployment. However, a full bonding equilibrium may be inefficient, relative to the first best, because the NEC_m may bind, yielding a corner solution in (13), because the SC may just bind, yielding a corner solution in (15), or the NEC_M may bind, also yielding a corner solution in (15). Let L^*_M be implicitly defined by $Mf'(L^*_M) = W_R$. Then full bonding equilibrium can be characterized as follows.

Proposition 1: In a full bonding equilibrium which satisfies the intuitive criterion and assumption 1

(i) either $L_m = L^*_m$ or $L_m < L^*_m$, the NEC_m is binding, and $B_m = 1/\theta$, and

(ii) either $L_M = L^*_M$ or $L_M < L^*_M$ in which case $B_M = 1/\theta$. In the latter case either the SC binds and $(M-m)f'(L_M) \leq 1/\theta$ or the NEC_M binds.

Proof: (i) follows immediately from examination of (13). From examination of (15) it follows that when its internal price is M the firm chooses its single period profit maximizing employment level, if both the SC and the NEC_M do not bind. When either the SC or the NEC_M binds, the firm does not utilize redundant bonds, i.e., $B_M = 1/\theta$. When the SC binds, it must be that an increase in employment, accompanied by the required wage increase to satisfy the SC, does not raise firm profit. This yields the condition, $(M-m)f'(L_M) < 1/\theta$. SS

Proposition 1 says that deviations from the single period profit maximizing solution occur because there is risk of bond expropriation, either by the firm when its internal price is m, directly through its own contract offer or indirectly through the SC by masquerading as if its

internal price were M, or by the firm when its internal price is M directly through its own contract offer. When a full bonding equilibrium exists where the SC binds, the condition that the firm finds it optimal to pay workers just their reservation wage when its internal price is M implies that it must be too costly for the firm to signal its type so that it can expand its employment. Note that in a separating equilibrium, deviations from single period profit maximization cannot occur in both states. However, it is possible to have a full bonding, pooling equilibrium where $L_m = L_M$, the NEC$_m$ is binding, the SC is binding, and it is too costly for the firm to signal its type so that it can raise employment when its internal price is M.

<u>IV.II Efficiency Wage Equilibrium</u>

When the SC binds in (15) and $W_M > W_R - F_M$, the firm pays its workers above their reservation wage to signal that its internal price is M. We refer to such an equilibrium as an efficiency wage equilibrium. Observe that this is more likely to occur the greater is the difference, $M - m$, and the greater is the quotient, M/m. We characterize efficiency wage equilibrium as follows.

<u>Proposition 2</u>: In an efficiency wage equilibrium which satisfies the intuitive criterion and assumption 1

(i) either $L_m = L_m^*$ or $L_m < L_m^*$, the NEC$_m$ binds, and $B_m = 1/\theta - F_m$; and

(ii) $L_M < L_M^*$, $Mf'(L_M) > W_M$, $B_M = 1/\theta - F_M$, and $mf'(L_M) + B_M \leq Mf'(L_M)$. If the NEC$_M$ doesn't bind then $Mf'(L_M) \leq mf'(L_M) + 1/\theta$

<u>Proof</u>: (i) follows immediately from examination of (13). Since the SC binds, it follows from (14) and (15) that $B_M = 1/\theta - F_M$. When the NEC$_M$ doesn't bind the first order condition for the optimal employment level in (15) is

(*) $mf'(L_M) + B_M \le Mf'(L_M) \le mf'(L_M) + 1/\theta$[23]

Assumption 1 can be rewritten as $B_M > W_M(M - m)/M$. Substituting this into the first inequality in (*) and subtracting W_M from both sides yields

(**) $\frac{m}{M}[Mf'(L_M) - W_M] < Mf'(L_M) - W_M$.

It follows that $Mf'(L_M) > W_M$, from which $L_M < L_M^*$ also follows. When the NEC$_M$ does bind the first order condition for the optimal employment level in (15) is just the first inequality in (*). The remaining argument is similar to the one just given. §§

When efficiency wage equilibrium exists and the NEC$_M$ doesn't bind, which we assume in the following discussion, there will be a continuum of such equilibria. See figure 4 for clarification on this point. We focus on the particular member in this class which yields the highest value of the firm, where the firm's isoprofit curve in state M is tangent to the SC from the left.[24] This member is depicted as point A in figure 4. From proposition 2 it follows that there is underemployment in such an equilibrium. We can now reconcile this underemployment result with the charging of up front fees when the firm is in the low productivity state. First, note that the firm cannot charge the up front fees in the high productivity state, because doing so would violate the SC. Thus, the best the firm can do to recoup the rents it pays its employees in the process of signaling is to charge fees in the low productivity state. Observe that all the rents are, in fact, recovered by the firm but the outcome is not as good as it would be were the firm able to charge the fees in the high productivity state. The difficulty is that the firm cannot credibly precommit to a higher employment level when in the high productivity state. When the firm pays efficiency wages in the high productivity state it, in effect, creates a positive externality for the firm in the low productivity state, because the entry fees are collected in the low productivity

[23] Observe that $B_M(L)L$ is not differentiable at $L = L_M$. Thus, there will be a kink in the SC at this employment level. Because of the kink in the SC there need not be a tangency between the SC and the relevant isoprofit curve in equilibrium.

[24] Since this tangency occurs at the kink in the SC, what we mean is that the isoprofit curve is tangent to the curve generated by the constraint (14), as written, ignoring the possibility of job rationing.

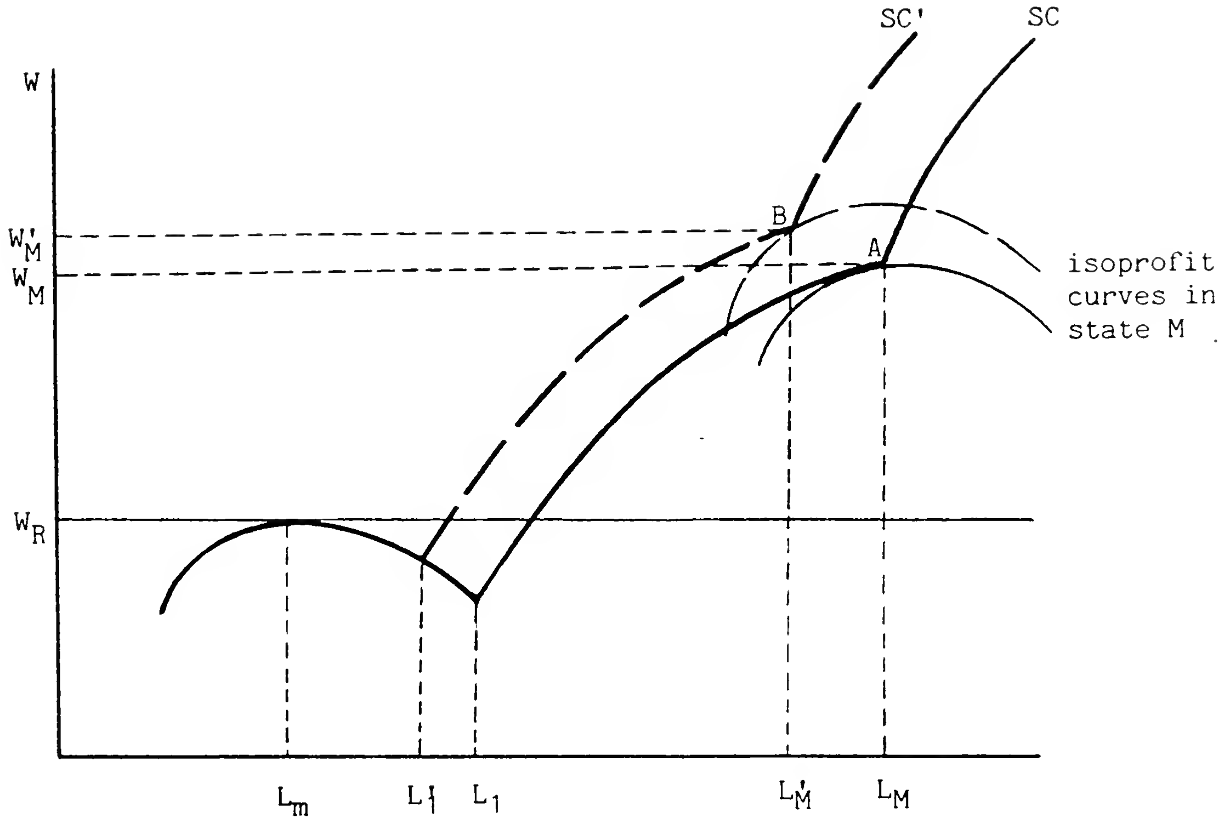

Figure 4

state and the magnitude of the entry fee depends on the size of the efficiency wage. Sequential rationality necessitates that the firm maximize current period profit when deciding which contract to offer, ruling out the possibility that the firm internalizes this apparent externality. This is why the equilibrium is inefficient.

This view of firm equilibrium differs from the standard efficiency wage equilibrium. The signaling aspect of equilibrium offers some very interesting hypotheses concerning wage determination. For example, larger scale firms, those with higher values of M, should pay higher wages, as should firms which have greater volatility in their value. Thus, this signaling approach provides an explanation for industry-wide wage differentials which does not require assuming that monitoring intensity varies across firms, the standard explanation of wage differentials provided by efficiency wage theory.[25]

IV.III Ruling Out Compensating Differentials in Pooling Equilibrium

The signaling equilibrium which we have demonstrated is not contingent on workers' initial beliefs, as long as workers are not certain that the firm's internal price is M. This result may appear puzzling. If workers are almost certain that the firm's internal price is M, why should the firm engage in costly signaling? That is, shouldn't a pooling equilibrium prevail in this case? In such a pooling equilibrium $\beta_l = \alpha_l$ and there will be a small amount of expropriation risk. The expected capital loss that a worker faces in such a pooling equilibrium is $\beta_l B_l$ and the effective performance bond, i.e., the value of the performance bond net of the expected capital loss, is $(1-\beta_l)B_l$. Let $F(\beta_l)$ denote the expected future rents per worker in such a pooling equilibrium,

$$(18)\quad F(\beta_l) = (1-\beta_l)(1-a)(W_M - W_R)/(r+a).^{26}$$

[25] See Katz (1986).

[26] Again, this assumes that employment in the pooling equilibrium does not exceed L_M.

Then $B_l = 1/[\theta(1-\beta_l)] - F(\beta_l)/(1-\beta_l)$ and $W_l = W_R - F(\beta_l) + \beta_l B_l$. Note that the contract wage contains a term which equals the expected capital loss due to bond expropriation. In the following proposition we show that such pooling equilibria are inconsistent with the intuitive criterion.

Proposition 3: If the equilibrium satisfies the intuitive criterion and $\alpha_l > 0$, then there does not exist a pooling equilibrium where the firm pays a compensating differential.

Proof: Were such an equilibrium to exist, then $B_l - W_l = 1/\theta - W_R$, independent of β_l. Thus, as long as employment stays constant, the firm has no incentive to affect β_l when its internal price equals m. However, W_l is increasing in β_l, so the firm has incentive to signal its internal price when this price equals M. In particular, if the firm were to offer an alternate contract where the bond was reduced by dB, the wage was reduced by $\beta_l\, dB$, and employment was slightly lowered, then the intuitive criterion implies that workers should be certain that the firm's internal price is M upon observing this alternate contract. Indeed, as long as workers are certain that the firm is in state M upon observing the alternate contract, it follows that the alternate contract has both a larger effective bond and a larger effective wage than the original contract. Thus, workers should both accept the alternate contract and put forth effort once the contract is accepted. The availability of such an alternate contract undermines the pooling equilibrium. §§

Note 1: When $\alpha_l = 0$ the argument in the proof of proposition 3 breaks down, since there is no need to pay a compensating differential in this case. Indeed, C^* is discontinuous at $\alpha_l = 0$.

Note 2: It might appear that there exists pooling equilibrium where there is no risk of bond expropriation but where the wage paid is less than W_R. In such an equilibrium, the firm finds that signaling to expand its employment is too costly, when in state M, and the firm finds it advantageous to pool to avoid paying a higher wage, when in state m. By essentially the same

argument as the one given in the proof of proposition 3, one can also rule out such pooling equilibria.

VI. Conclusion

In the body of the paper we treated the firm's monitoring intensity as exogenous. We now briefly consider the implications of the firm choosing its level of monitoring. Standard theoretical argument suggests that the optimal level of monitoring is essentially zero, since monitoring requires real resource costs and incentives can be maintained by simply increasing the size of the punishment when detection occurs. This argument ignores the issue of how societal norms restrict the size of punishments.[27] Our paper suggests another problem which mitigates against the use of arbitrarily large punishments. These punishments cannot be made credible without the firm deviating from its first best employment levels. This suggests that there is a tradeoff between monitoring costs, on the one hand, and allocational costs via underemployment, on the other. From the perspective of positive theory, this tradeoff implies that optimal contracting equilibrium does not yield first best employment levels. This follows since the marginal monitoring cost is positive while the marginal cost to altering the labor allocation from its first best level is essentially zero.

Before closing, we turn to one last point concerning interpretation. In equilibrium of our model, the firm never expropriates the performance bonds. We believe that expropriation may occur in actuality. This leads to the question of how expropriation might be manifest. We have modeled the firm moral hazard under the assumption that the firm goes out of business once it has expropriated performance bonds and, consequently, our model may suggest to the reader that expropriation occurs via the firm declaring bankruptcy. This interpretation is valid as long as it is the firm, rather than the firm's creditors, which receives the proceeds from bond

---

[27] See note 4.

expropriation. However, we did not intend for the interpretation of our model to be restricted to the case where the firm declares bankruptcy. Indeed, in light of the recent rage of takeovers apparently motivated by substantial returns from labor bashing and, assuming that most bonding in actuality is implicit and occurs as deferred wages, our preferred interpretation is that expropriation occurs when new management, less bound by implicit commitments with employees than was the previous management, bargains down wages by utilizing the threat of massive layoffs in the event that the wage bill is not substantially reduced.[28] Whether this phenomenon is actually an example of bond expropriation or merely the predictable outcome which results when diligent management succeeds turf protecting management is moot. Certainly, the issue is sufficiently provocative to warrant further research.

[28] Schleifer and Summers (1987) make essentially the same point.

References

Abraham, K.G., and Farber, H.S., 1987, "Job Duration, Seniority, and Earnings," American Economic Review, 77,3: 278-297.

Akerlof, G.A., and Katz, L.F., 1986, "Do Deferred Wages Dominate Involuntary Unemployment as a Worker Discipline Device?" National Bureau of Economic Research, Working Paper No. 2025.

Akerlof, G.A., and Yellen, J., 1986, "Introduction," in Akerlof and Yellen, eds., Efficiency Wage Models of the Labor Market, Cambridge University Press.

Arvan, L., and Esfahani, H.S., 1987, "Performance Bonds, Firm Reputation, and Free-Entry Equilibrium," University of Illinois BEBR Working Paper No. 1387.

Bulow, J.I., and Summers, L.H., 1986,"A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment," Journal of Labor Economics, 1,3: 376-414.

Carmichael, L.H., 1985, "Can Unemployment Be Involuntary? Comment," American Economic Review, 75,5: 1213-1214.

Cho, I.K., and Kreps, D.M., 1987,"Signaling Games and Stable Equilibria," Quarterly Journal of Economics, 102,2: 179-222.

Dickens, W., Katz, L.F., Lang,K., and Summers,L.H., 1987, "Employee Crime, Monitoring, and the Efficiency Wage Hypothesis," Mimeo.

Eaton, B.C., and White,W.,1982, "Agent Compensation and Limits of Bonding," Economic Inquiry, 20,3: 330-343.

Hart, O., 1983, "Optimal Labor Contracts under Asymmetric Information," Review of Economic Studies, 50: 3-36.

Katz, L.F., 1986, "Efficiency Wage Theories: A Partial Evaluation," National Bureau of Economic Research, Working Paper No. 1906.

Kreps, D.M., and Wilson, R. 1982, "Sequential Equilibria," Econometrica, 50: 863-894.

Malcomson, J., 1984 "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92,3: 486-507.

Medoff, J., and Abraham, K., 1980, "Experience, Performance and Earnings," Quarterly Journal of Economics, 95: 703-736.

Milgrom, P., and Roberts, J., 1986, "Price and Advertising Signals of Product Quality," Journal of Political Economy, 94,4: 796-821.

Quarterly Journal of Economics, Supplement, 1983.

Schleifer, A., and Summers, L.H., 1987, "Breach of Trust in Hostile Takeover" NBER Working Paper, No. 2342, August 1987

Shapiro, C., and Stiglitz, J.E., 1984, "Equilibrium Unemployment as a Worker Discipline Device," American Economic Review, 74,4 433-444

Shapiro, C. and Stiglitz, J.E., 1985, "Can Unemployment Be Involuntary? Reply" American Economic Review, 75.5: 1215-1217

Yellen, J., 1984, "Efficiency Wage Models of Unemployment," American Economic Review: Papers and Proceedings, 74.2. 200-205

Papers in the Political Economy of Institutions Series

No. 1 Susan I. Cohen. "Pareto Optimality and Bidding for Contracts." Working Paper # 1411

No. 2 Jan K. Brueckner and Kangoh Lee. "Spatially-Limited Altruism, Mixed Clubs, and Local Income Redistribution." Working Paper #1406

No. 3 George E. Monahan and Vijay K. Vemuri. "Monotonicity of Second-Best Optimal Contracts." Working Paper #1417

No. 4 Charles D. Kolstad, Gary V. Johnson, and Thomas S. Ulen. "*Ex Post* Liability for Harm vs. *Ex* *Ante* Safety Regulation: Substitutes or Complements?" Working Paper #1419

No. 5 Lanny Arvan and Hadi S. Esfahani. "A Model of Efficiency Wages as a Signal of Firm Value." Working Paper #1424

No. 6 Kalyan Chatterjee and Larry Samuelson. "Perfect Equilibria in Simultaneous-Offers Bargaining." Working Paper #1425